

87 GYDPYDFDLGEYNQKGTETRFSGKQELINMINTAHAYGIV 146
 122 PFVNDYTWTFDSKVASGKYTANYLDHPNHLHAGDSGTGGYDPDI CHDKSWDOYWLWASQ 181
 147 PFVGDYTWTFDSKVASGKYTANYLDHPNHLHAGDSGTGGYDPDI CHDKSWDOYWLWASQ 206
 182 ESYAAYLRSIGIDAWRFDYVKGYPAPVVKWLNWGMGWA VGEYWDNVDVAVLNWAYSSGA 241
 207 ESYAAYLRSIGIDAWRFDYVKGYPAPVVKWLNWGMGWA VGEYWDNVDVAVLNWAYSSGA 266
 242 KVFDFALYKXDEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPAY 301
 267 KVFDFALYKXDEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPAY 326
 302 AFILTYEQOPTIFRDYEELNKKLNLIWIHNLHAGDSGTGGYDPDI CHDKSWDOYWLWASQ 361
 327 AFILTYEQOPTIFRDYEELNKKLNLIWIHNLHAGDSGTGGYDPDI CHDKSWDOYWLWASQ 386
 362 KPGILTYINLSSKAGRWVYVVKPAGACIHEYTGNLGGWVDKTVYSSGHWYLEAPAYDPA 421
 387 KPGILTYINLSSKAGRWVYVVKPAGACIHEYTGNLGGWVDKTVYSSGHWYLEAPAYDPA 446
 422 NGQYGSVMSYCGVG 436
 447 NGQYGSVMSYCGVG 461

RESULT 4

O33476
 ID O33476 PRELIMINARY; PRT; 461 AA.
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Alpha-amylose precursor.
 GN APKA.
 OS Pyrococcus kodakaraensis.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Thermococcus.
 OX NCBI_TaxID=69014;
 RN [1]_TaxID=69014;
 RP SEQUENCE FROM N.A.
 RC STRAIN=KOD1;
 RA Tachibana Y., Mendez L., Fujiwara S., Takagi M., Imanaka T.;
 RT "Cloning and expression of the alpha-amylose gene from the
 RT hyperthermophilic archaeon Pyrococcus sp. KOD1, and characterization
 of the enzyme."
 RL J. Ferment. Bioceng. 82:224-232 (1996).
 DR EMBL; D83793; BAA21130.1; -
 DR HSSP; P06278; 1VUS.
 DR GO; GO:0004556; F:alpha-amylose activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR006047; Alpha_amy1_cat.
 DR InterPro; IPR006046; Glyco_hydro_13.
 DR Pfam; PF00128; alpha-amylose; 1.
 DR PRINTS; PR00110; ALPHAAMYLASE.
 KW Signal.
 FT SIGNAL.
 FT CHAIN 1 26 POTENTIAL.
 FT CHAIN 27 461 ALPHA-AMYLASE.
 SQ SEQUENCE 461 AA; 52213 MW; FCC131A93DC03123 CRC64;

Query Match 91.4%; Score 2248; DB 1; Length 461;
 Best Local Similarity 90.6%; Pred. No. 5.9e-144;
 Matches 394; Conservative 18; Mismatches 23; Indels 0; Gaps 0;
 2 AKYSELEKGVINQAFYWDVPSGGIWDTIROKIPENTWDAGISAINIPPASKNGGAYSM 61
 27 AKYSELEKGVINQAFYWDVPSGGIWDTIROKIPENTWDAGISAINIPPASKNGGAYSM 86
 62 GYDPYDFDLGEYNQKGTETRFSGKQELINMINTAHAYGIVINHRAGDLEWN 121
 87 GYDPYDFDLGEYNQKGTETRFSGKQELINMINTAHAYGIVINHRAGDLEWN 146

122 PFVNDYTWTFDSKVASGKYTANYLDHPNHLHAGDSGTGGYDPDI CHDKSWDOYWLWASQ 181
 147 PFVGDYTWTFDSKVASGKYTANYLDHPNHLHAGDSGTGGYDPDI CHDKSWDOYWLWASQ 206
 182 ESYAAYLRSIGIDAWRFDYVKGYPAPVVKWLNWGMGWA VGEYWDNVDVAVLNWAYSSGA 241
 207 ESYAAYLRSIGIDAWRFDYVKGYPAPVVKWLNWGMGWA VGEYWDNVDVAVLNWAYSSGA 266
 242 KVFDFALYKXDEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPAY 301
 267 KVFDFALYKXDEAFDNKNIIPALVSALONGQTVVSRDPFKAVTFVANHDTDIINWKYPAY 326
 302 AFILTYEQOPTIFRDYEELNKKLNLIWIHNLHAGDSGTGGYDPDI CHDKSWDOYWLWASQ 361
 327 AFILTYEQOPTIFRDYEELNKKLNLIWIHNLHAGDSGTGGYDPDI CHDKSWDOYWLWASQ 386
 362 KPGILTYINLSSKAGRWVYVVKPAGACIHEYTGNLGGWVDKTVYSSGHWYLEAPAYDPA 421
 387 KPGILTYINLSSKAGRWVYVVKPAGACIHEYTGNLGGWVDKTVYSSGHWYLEAPAYDPA 446
 422 NGQYGSVMSYCGVG 436
 447 NGQYGSVMSYCGVG 461

pdate)
 (update)

coccales; Thermococcaceae;

11 and Molecular
ims-Champagne-Ardenne,
51687, France

lis"

FEAT

PAKAETLENGVIMQAFYWDVPG
/SMGYDPYDFDLGEYKQGSVE
JLEWNPFTNSYATWTFKSVASGK
HLMASNESYAYLPSIGIDAWR
SWAYDSGAKUPDFPLYKMDER
XLIWNKYPAYAFILITYEGQPAI
IDELIFVRENGYDKFGLITYINL
USGRVYLEAPADHPANGQYGSV

WSYCVGVG"

ORIGIN

Query Match 76.5%; Score 1003.4; DB 1; Length 2705;

Best Local Similarity 86.0%; Pred. No. 3.4e-216;

Matches 1112; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY	19	CTGGAAAAGGCGGGTCAATGATGAGCGGTCTTACTGGGAGCGTGCCTTCAGGAGGAATA	78
DB	505	CTTGAGAACGGCGGTCTAATGAGCGCTTCTACTGGGAGCGTCCCGAGTGGAGGATC	564
QY	79	TGGTGGGACACAATACGGCAGAGATACGGAGTGGTACGATGCGGGAATCTCCGCAATA	138
DB	565	TGGTGGGACACCATAGCCCGAGAGATACCGGACTGGGCGAGCGCGGATTCGCGCAATA	624
QY	139	TGGATTCCCCCGGAGGAGGCGATGGCGCGCTATTGATGGGCTAGACCCCTAC	198
DB	625	TGATTTCCTCCCGGAGTAAAGGCGATGAGCGCGCTATTGATGGGCTAGACCCCTAC	684
QY	199	GACTTCTTTGACTCGGTGAGTACGACAGAGGAGACGGTAGACACGCGCTTTGGCTCC	258
DB	685	GATTTCTTCGACTCGGTGAGTACTACGAGAGGAGGAGGTTGAGACCGGCTTCGGATCA	744
QY	259	AAGCAGGAGCTGTGAGCATGATTAACACCGCCGCGCTATGCGTGAAGTATAGCC	318
DB	745	AAGAGGAGCTGTGAGCATGATTAACACCGCCGCGCTATGCGTGAAGTATAGCC	804
QY	319	GATATAGTCATCAACACCGCGCGCGCTGAGTGGAGTGAACCCCTTCGTGAACGAC	378
DB	805	GACATAGTCATCAACACCGCGCGCGCTGAGTGGAGTGAACCCCTTCGTGAACGAC	864
QY	379	TATACCTGACGACCTTCTCAAGGTGCGGTGGGTAAATACAGGCGCACTACCTCGAC	438
DB	865	TACACCTGACGACCTTCTCAAGGTGCGGTGGGTAAATACAGGCGCACTACCTCGAC	924
QY	439	TTCCACCCGAAACGAGCTCCATGCGGCGGATTCGGAACATTTGGAGGCTATCCGACATA	498
DB	925	TTCCACCCGAAACGAGCTTCACGCGGCGGATTCGGAACATTTGGAGGCTATCCGACATA	984
QY	499	TGCCACGACAGAGCTGGGACCAAGTACTGGCTCTGGGCGAGCGAGAGCTACGCGCA	558
DB	985	TGCCACGACAGAGCTGGGACCAAGTACTGGCTCTGGGCGAGCAACGAAAGCTACGCGCC	1044
QY	559	TATCTCAGGAGCATCGGCATCGATGCTCGGCGCTTCGACTACGTCAAGGGCTATGCTCC	618

DB	1045	TACCTCCGAGCATCGGCATCGACGCGTGGCGTTCGACTACGTCAAGGGCTACGCTCCC	1104
QY	619	TGGGTCTGCTCAAGGACTCGCTGAACCTGCTGGGAGGCTGGGCGGTGGAGAGTACTGGGAC	678
DB	1105	TGGGTCTGCTTAAGAACTGGCTGAACCGGTGGGCGGCTGGGCGGTGGAGAGTACTGGGAC	1164
QY	679	ACCAACGTCGACGCTGTTCTCAACTGGGCTATCTCGAGCGGTGCCAACAGGTCTTTGACTTC	738
DB	1165	ACCAACGTCGACGCTCCTGAGCTGGGCTACGACAGCGGTGCTAAAGTCTTCGACTTC	1224
QY	739	GCCCTCTACTACAGATGGATGGGCTTTGACACCAAAACATTCCTCAGCGCTCGTCTCT	798
DB	1225	CGCTCTACTACAGATGGAGGCGCTTCGATAACCAACATCTCCCGGCTCGTGGAC	1284
QY	799	GCCCTTCAGAACCGGCGACGCTGTTGCTCCCGGACCGCTTCAAGGCGGTAAACCTTTGTA	858
DB	1285	GCCCTCAAGAACGGGCGACGCTGCTCAGCGCGGACCGTTCRAAGCGGTGACCTTCGT	1344
QY	859	GCAACACGACGACCGGATATTAATCTGGAAACAGTATCAGGCTACGGTTCATCTCACC	918
DB	1345	GCAACACGACGATACCAACATTAATCTGGAACAAAGTATCGGCTACGCTTCATCTCACC	1404
QY	919	TACGAGGCGGACCGGACATATTTCTACCGGACTACGAGGAGTGGCTCAACAGGATAAG	978
DB	1405	TATGAGGACGACCGGCAATTTCTACCGGACTACGAGGAGTGGCTCAACAGGACGAG	1464
QY	979	CTCAAGAACCTCATCTGATATCATGAGAACCTCGCGGAGGAGGACGACATAGTCTAC	1038
DB	1465	CTCAGGAACCTCATCTGATATCATGAGAACCTCGCGGAGGAGGACGACATAGTCTAC	1524
QY	1039	TACGATAACGATGATCACTCTTCTGTCAGGAACGCTACGCGGACCAAGCGGGCTTATA	1098
DB	1525	TACGACGAGGAGGAGCTTATCTTCTGTCAGAACGCTACGCGGACCAAGCGGGACTGATA	1584
QY	1099	ACTTACATCAACTAGGCTCGAGCAAGCGCGGAGGTTTATGTCCGAAAGTTCGCG	1158
DB	1585	ACTTACATCAACTCGGCTCAGCAAGCGCGGAGGTTTATGTCCGAAAGTTCGCG	1644
QY	1159	GCGGCTGATCCAGAGTATCTGTTAACTCGGAGGCTGGGTAGACAGTACGCTTAC	1218
DB	1645	GCGTCTGTCATACAGAGTACACCGGCAACCTCGCGGCTGGATTGACAGTGGGTTAC	1704
QY	1219	TCAGCGGCTGGTCTTCTCTGAACTCCAGCTTACGACCTCGCAACCGGCGGAGTATGCG	1278
DB	1705	TCAGCGGCTGGTCTTCTCTGAACTCCAGCTTACGACCTCGCAACCGGCGGAGTATGCG	1764
QY	1279	TACTCGTGGAGTACTGCGGCTGGGCTGA	1311
DB	1765	TACTCGTGGAGTACTGCGGCTGGGCTGA	1797

RESULT 6

DB3793

LOCUS

DB3793

DEFINITION

Pyrococcus sp. DNA for alpha-amylase, complete cds.

ACCESSION

DB3793

VERSION

DB3793.1 GI:2251107

KEYWORDS

Amk; alpha-amylase.

SOURCE

Pyrococcus sp.

ORGANISM

Pyrococcus sp.

Archaea; Euryarchaeota; Thermococci; Thermococcales;

Thermococcaceae; Pyrococcus.

REFERENCE

1 (sites)

Tachibana, Y., Mendez, L., Fujiwara, S., Takagi, M. and Imanaka, T.

Cloning and expression of the alpha-amylase gene from the

hyperthermophilic archaeon Pyrococcus sp. KOD1, and

characterization of the enzyme

2 (bases 1 to 2179)

J. Ferment. Bioeng. 82, 224-232 (1996)

AUTHORS

Tachibana, Y., Mendez, L., Takagi, M. and Imanaka, T.

Direct Submission

JOURNAL

Submitted (05-MAR-1996) Yoshihisa Tachibana, Osaka University,

2179 bp DNA linear BCT 01-FEB-2000

Faculty of Engineering, Biotechnology; 2-1, Yamadaoka, Suita, Osaka
565, Japan (Tel:06-879-7442, Fax:06-879-7448)

Location/Qualifiers
1. .2179

/organism="Pyrococcus sp."
/mol_type="genomic DNA"

/strain="KOD1"
/db_xref="taxon:33866"

452.11837
/gene="Apka"

452.11837
/gene="Apka"

/codon_start=1
/transl_table=11

/product="alpha-amylase"
/protein_id="BA21130.1"

/db_xref="GI:2251108"
/translation="MKKFLVALLTMEFVSMVAAPASAAKYSELEGGVIMQAFYV

DVPGGIWMTIRSKIPEWYEAGISAIWIPPAKNGMGAYSGYDYPDFDLGRYNQK
GTVEFRGKDELINMINTAHAYGIKVIADIVINHRAGDLEWNPFGDYTWDFSKV
ASGYTANYDHFENEVKCDEGTFGGFPDIAHEKEDQHLWASDESYAYLRSIGV
DAWFDYVKGAWVWDLNMGWAGVGEVDNVDALLNWAYSSGAKVDFPLIYK
MDEAFDNTNI PALVDALONGSTVVSRRDPKAVTFVANHTDIIWKYPAVAFILYEG
QPVLFYRDEMLKDLNLIHDEHLAGGSTIVYDSDLELIFVRNGDSKRLGIT
YINLGSSKGVWVVPKFAAGACIHEVTGNLGNVDKXVSSGNVLEAPATPDPSGQY
GTVWSVCGVG"

sig_peptide
452.1529

mat_peptide
530.11834

/gene="Apka"
/product="alpha-amylase"

ORIGIN

Query Match 75.6%; Score 991.2; DB 1; Length 2179;
Best Local Similarity 84.9%; Pred. No. 26-213;
Matches 1110; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

2Y 4 GCACAGTACTCGAGCTCGAAAGCGGGGTCTAATATGAGGCGTTCTACTGGGACGTG 63
Db 530 GCACAGTACTCGAGCTCGAAAGCGGGGTCTAATATGAGGCGTTCTACTGGGACGT 589
2Y 64 CTTTCAGGAGGAATATGTTGGGACACATACCGGAGGATACCGGAGTGGTACGATGCC 123
Db 590 CCAGCGGAGGATCTGTTGGGATACATCAGAGGAGATACCGGAGTGGTACGAGGT 649
2Y 124 GGAATCTCGCAATATGATATCCCGCGGAGGAGGAGGATGCGGCGCGCTATTTCGATG 183
Db 650 GGAATCTCGCAATCTGGAATCCCGCGGAGGAGGAGGATGCGGCGCGCTATTTCGATG 709
2Y 184 GGCTACGACCCCTACGACTTCTTTGACCTCTGAGTACGACACGAGAGGAGGAGGAG 243
Db 710 GGCTACGACCCCTACGACTTCTTTGACCTCTGAGTACGACACGAGAGGAGGAG 769
2Y 244 ACSCGCTTTGGCTCCAGCAGGAGCTCGTACATGATTAACACCGCCGCGCTATGGC 303
Db 770 ACTCGCTTTGGCTCCAGCAGGAGCTTATACATGATTAACACCGCCGCGCTATGGC 829
2Y 304 ATGAGGTAATAGCGGATATAGTCTCAACACCGCGCGCGGCGGCTGAGCTGGAGTGAAC 363
Db 830 ATGAGGTAATAGCTCATATGCTCAACACCGCGCGCGGCGGCGGAGACCTCGAGTGGAC 899
2Y 364 CCCTTCGAGACGACTATACCTGGACCGGCTTCTCAAGAGTGGCTGGGTTAATACAG 423
Db 890 CCCTTCGAGACGACTATACCTGGACCGGCTTCTCAAGAGTGGCTGGGTTAATACAG 949
2Y 424 GCCAACTACCTCGACTTCCACCGGACGAGCTCCATGCGGGCGGATTCGCGAACATTTGGA 483
Db 950 GCCAACTACCTCGACTTCCACCGGACGAGCTCCATGCGGGCGGATTCGCGAACATTTGGA 1009
2Y 484 GGCTATCCGACATATCCACGACGAGAGCTGGGCGGAGCTGCTGCTCTGGGCGGCGAG 543
Db 1010 GGCTATCCGACATATCCACGACGAGAGCTGGGCGGAGCTGCTGCTCTGGGCGGCGAG 1069

QY 544 GAGAGCTACCGCGCATATCTCAGAGAGATCGCATCGATGCTGGCGTTCGACTAGCTC 603
Db 1070 GAGAGCTACCGCGCATATCTCAGAGAGATCGCATCGATGCTGGCGTTCGACTAGCTC 1129
QY 604 AAGGCTATGCTCCCTGGGTCGTCAGGAGTGGCTGAACTGGTGGGAGGCTGGGCGGTT 663
Db 1130 AAGGCTATGCTCCCTGGGTCGTCAGGAGTGGCTGAACTGGTGGGAGGCTGGGCGGTT 1189
QY 664 GAGAGTACTGGGACACCAACGTCGAGCGTCTTCACTGGGCGATATCTGAGCGGTGCC 723
Db 1190 GGTGAGTACTGGGACACCAACGTCGAGCGTCTTCACTGGGCGATATCTGAGCGGTGCC 1249
QY 724 AAGGCTTCTGACTTCGCGCTCTACTACAGATGAGTGGGCGCTTTGACAAACAAACATT 783
Db 1250 AAGGCTTCTGACTTCGCGCTCTACTACAGATGAGTGGGCGCTTTGACAAACAAACATT 1309
QY 784 CCAGGCTCGTCTCTGCGCTTCAAGCGGCGAGATGTTGTTCTCCCGCGACCGGTTCAAG 843
Db 1310 CCAGGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1369
QY 844 GCGGTAACTTTGTAGCAAAACCAACGACGACGATATATCTGGAACAAAGTATCCAGCTTAC 903
Db 1370 GCGGTAACTTTGTAGCAAAACCAACGACGACGATATATCTGGAACAAAGTATCCAGCTTAC 1429
QY 904 GCGTTCATCTCTCCTAGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 963
Db 1430 GCGTTCATCTCTCCTAGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1489
QY 964 CTCACAAAGGATAAGCTCAAGAACCTCATCTGGATACATGAGAACCTCCCGGAGGAGGAG 1023
Db 1490 CTCACAAAGGATAAGCTTAAACAACTAATCTGGATACACGACCACTCCCGGAGGAGGAG 1549
QY 1024 ACCGACATAGTCTACTACGATAACGATGAATCTCTCTCAGGAGGAGGAGGAGGAGGAG 1083
Db 1550 ACCGACATAGTCTACTACGATAACGATGAATCTCTCTCAGGAGGAGGAGGAGGAGGAG 1609
QY 1084 AAGCGGCGGCTTATTAACCTATCACTAGGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1143
Db 1610 AAGCGGCGGCTTATTAACCTATCACTAGGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1669
QY 1144 GTGCGGAGGTTCCGCGGCGGCTGATCCAGAGGATATCTGTTAACTCTCGGAGGAGGAGGAG 1203
Db 1670 GTGCGGAGGTTCCGCGGCGGCTGATCCAGAGGATATCTGTTAACTCTCGGAGGAGGAGGAG 1729
QY 1204 GACAGTACTGCTACTCAGCGGCTGGGCTCTATCTCGAGGCTCCAGCTTACGAGGAGGAG 1263
Db 1730 GACAGTACTGCTGAGTGGAGGCGCTGGGCTCTATCTCGAGGCTCCAGCTTACGAGGAGGAG 1789
QY 1264 AACGGGCGAGTATGGCTACTCCGCTGTGGAGCTACTGCGGGGTTGGGCTGA 1311
Db 1790 AACGGGCGAGTATGGGCTACTACCGCTCTGGAGCTACTGCGGGGTTGGGATGA 1837